



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Embodiment, Consciousness, and Neurophenomenology

Citation for published version:

Rupert, R 2015, 'Embodiment, Consciousness, and Neurophenomenology: Embodied Cognitive Science Puts the (First) Person in Its Place', *Journal of consciousness studies*, vol. 22, no. 3-4, pp. 148-180.
<<http://www.ingentaconnect.com/content/imp/jcs/2015/00000022/F0020003/art00009>>

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Journal of consciousness studies

Publisher Rights Statement:

© Rupert, R. (2015). Embodiment, Consciousness, and Neurophenomenology: Embodied Cognitive Science Puts the (First) Person in Its Place. *Journal of consciousness studies*, 22(3-4), 148-180.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Embodiment, Consciousness, and Neurophenomenology: Embodied Cognitive Science Puts the (First) Person in Its Place

Forthcoming in a special issue of *Journal of Consciousness Studies*

Robert D. Rupert
University of Colorado, Boulder, and University of Edinburgh

Introduction

Talk of embodiment has permeated discourse in cognitive science and well beyond, extending from its contemporary roots in robotics (Brooks 1999) to discussions of language comprehension (Kaschak et al. 2005), technological augmentation (Clark 2003), memory (both at the level of the individual – Glenberg 1997 – and the group – Barnier et al. 2008), Elizabethan performance (Tribble 2005), mathematical reasoning (Lakoff and Nunez 2000), religious experience (Barsalou et al, 2005), everyday communication (Goldin-Meadow 2003), concepts (Barsalou 1999), vision-based problem solving (Ballard et al. 1997), and further still – for overviews, see Gibbs (2006) and Clark (2008). This blossoming embodiment-related literature catalogues, in fascinating detail, ways in which bodily motion, nonneural bodily structures, and environmental props ground, contribute to, and perhaps even partly constitute human cognitive processing. Much of this work manifests grand theoretical ambition as well. Sweeping generalization and revolutionary pronouncement make regular appearances, in the form of such provocative theses as that all cognition is body-based, that there are no abstract concepts, that computational conceptions of cognition are bankrupt, and that functionalism is false.

In this paper, I attempt to sort out one of the most striking theoretical claims associated with embodied approach to cognition and mind: that an embodied cognitive

science accounts for consciousness.¹ This is no minor matter. If embodied theorists can make good on this promise, they will have achieved a monumental victory; they will have proven wrong the many philosophers and cognitive scientists who take the scientific understanding of consciousness to run hopelessly aground on the so-called hard problem of consciousness (Chalmers 1996), the problem of accounting for the “what it’s like” of conscious experience. According to Chalmers, any merely causal or functional story about physical processing misses the point. Such a story explains, at best, only what consciousness does or what causes it, without capturing what it’s like to be in a conscious state, the intrinsic nature of the experience of smelling a rose, for example, or of tasting an orange. Causal or functional accounts fail to bridge the explanatory gap between brute neural processes and the functional-role facts they determine, on the one hand, and the rich world of inner experience available to every conscious human, on the other (Levine 1983).

Many fans of embodiment accept Chalmers’s challenge as issued, or at least seem to, and argue, in response, that careful attention to bodily processes can explain conscious experience without residue. So far as I can tell, this tack makes no progress on the hard problem (Rupert 2006). Moreover, for independent reasons, I think it represents a fundamental mistake in our understanding of the relation between consciousness and embodiment-oriented research in cognitive science. In my view, embodiment-based cognitive science – on balance and in conjunction with empirical research that is not particularly connected to the embodied movement – marginalizes the intuitions that generate the hard problem (Rupert 2011). Bodily-based factors participate actively in the

¹ For attempts to sort out some other theoretical claims associated with the embodied perspective, see Rupert (2006; 2009, Ch. 11) and Shapiro (2010).

subconscious processing that produces the behavior we associate with consciousness, attention, awareness, and intentional control. *This* – the subconscious, partly bodily-based contribution to the behavior associated with conscious awareness – is what much of the most impressive empirical work on embodiment probes. Importantly, the combination of extant embodied and not-so-embodied research also helps to identify the genesis of erroneous intuitions about the “what it’s like” of consciousness and about the intrinsic qualitative characters supposed to be possessed by our conscious experiences. Thus, as I see things, an embodied cognitive science constitutes part of an error theory, that is, an explaining away of the intuitions that give rise to the supposed hard problem.²

My take on embodied cognitive science stands at odds not only with the views of embodiment theorists who aim to solve the hard problem head on, but also with the views of embodiment theorists who dismiss the idea that a properly scientific approach to consciousness should concern itself with the hard problem and instead attempt to incorporate unreduced first-personal phenomenal experience directly into a reconceived cognitive science (Varela 1996). Such attempts presuppose the veridicality and utility of first-person intuitions about consciousness – or of their refined counterparts yielded by phenomenological reduction. I argue that successful embodiment-based research and the way in which it contributes to our understanding of human cognition undermines this neurophenomenology (as it’s often called), at least as a cognitive-scientific standard-bearer for the embodied approach to consciousness.

² The view I express here and return to later in the paper is not necessarily eliminativist about conscious experience. Rather, I claim that intuitions about the theoretically important properties of conscious experience have misled many philosophers. This is not to say that eliminativism about the natural kind *conscious experience* is off the table. It may be that our talk of conscious experiences tracks a number of genuine states that do not share a cluster of theoretically important properties, in which case ‘conscious experience’ as a kind term may refer only to a nominal kind (cf. Machery 2009 about concepts, Rupert 2013 about cognition).

Here is how the essay unfolds. Section One distinguishes between and elaborates upon various versions of the embodiment theorist's ambitious thesis, that an embodied cognitive science solves the hard problem as posed. Section Two criticizes these embodiment-related approaches to the hard problem; such proposals either (a) help themselves to a certain amount of unexplained phenomenal consciousness (at the "bodily level") or (b) offer what amounts to more causal-functional modeling of physical processes, of the sort that, to the mind of someone who has accepted Chalmers's challenge as issued, can't possibly account for the intrinsic qualitative character of conscious experience. Section Three explores an alternative approach to the embodiment-oriented science of consciousness: neurophenomenology (Varela 1996). On this view, standard cognitive-scientific approaches cannot sort out the problem of phenomenal consciousness because they ignore the first-person perspective; and contrapositively, by incorporating a first-person method, cognitive neuroscience can solve the hard problem – by, in some sense, dissolving it. Section Four takes three-fold exception to this program. First, some versions of neurophenomenology take unreduced, phenomenally conscious experiences for granted, and thus amount to anti-scientific capitulation; they could be right as such, but they're not part of an embodied cognitive science, as science has been widely understood and practiced heretofore. Second, as it's been successfully pursued in an experimental context, the neurophenomenological method is standard cognitive science and not a manifestation of any distinctively embodied cognitive science – it recommends treating verbal reports as a legitimate stream of data, which is standard practice (and thus neither a solution to nor a dissolving of the hard problem). And third,

embodied cognitive science itself tends to marginalize the neurophenomenological method, a point pursued in the final section.

Section Five lays out an alternative gloss of embodied cognitive science and its relation to phenomenal consciousness. Embodied processes contribute frequently and powerfully to the production of the observable phenomena associated with consciousness and awareness. They do so via their participation in the swarm of activity occurring at what is often described as the subpersonal level. Part of what embodied cognitive science reveals, however, is that the so-called subpersonal level is the one and only psychological level. Thus, embodied cognitive science helps to show why the conversation surrounding the hard problem rests largely on a mistake, and it does so by casting more appropriately the first-person perspective that creates said problems – as a bit player that provides one source of data to cognitive science on a par with other forms of data (Dennett 1991, Piccinini 2010).

I. Embodiment and the Hard Problem

What have embodiment-oriented theorists had to say about consciousness? Here's a representative sampling:

From George Lakoff and Mark Johnson:

If you are a normal human being, you inevitably acquire an enormous range of primary metaphors just by going about the world constantly moving and perceiving. Whenever a domain of subjective experience or judgment is coactivated regularly with a sensorimotor domain, permanent neural connections are established via synaptic weight changes. Those connections, which you have unconsciously formed by the thousands, provide inferential structure and qualitative experience activated in the sensorimotor system to the subjective domains they are associated with. (1999, 57)

Along similar lines, Raymond Gibbs says, “Most generally, this examination of metaphor and linguistic action reveals how people use their intuitive phenomenological sense of their bodies to make sense of, and structure, more abstract conceptual domains” (Gibbs 2006, 104).

Gibbs emphasizes other aspects of the relation between consciousness and the body – particularly to do with movement, timing, and interaction with the environment – as well: “[U]nderstanding how consciousness arises from animate motion, and more specifically the dynamical interactions of brains, bodies in motion, and the environment, is the best way to close this explanatory gap and ultimately solve the ‘hard problem’” (2006, 263), and, “Recognizing the tight relationship between kinesthetic action and cognition (i.e., the decisions we make) demonstrates that there is not a mysterious gap between cognitive abilities and consciousness” (*ibid.*, 266).

Tony Chemero claims that adherents of his radically embodied approach to cognitive science can “explain experience in such a way that qualia [the states or properties that are the ‘what’s its like’] do not come up” (2009, 198). According to Chemero, the subject’s interaction with the environment is driven by and is partly constituted by affordances, roughly speaking, the relations that hold between the animal and its environment such that the animal can do certain things in that environment; having an experience is, in some sense, nothing more than having one’s behavior driven by such affordances: “Conscious experiences, that is, are what happen when animals pick up information about affordances” (*ibid.*, 201).

And there’s much more, some of which will be canvassed below. But, we can see from this initial sampling that embodiment-oriented authors don’t shy away from

questions about consciousness and the hard problem. What, however, *is* the embodied strategy for solving the hard problem? Is there a single one? In what follows, I briefly describe three strategies that might be derived from the work of the authors quoted above and much else in the embodied literature.

1. *Conscious bodily experience is basic; the remainder of conscious experience is constructed from this basic stock.*

This approach can be found in Lakoff and Johnson's work and in the first of the quotations from Gibbs given above. This general strategy begins by identifying simple and apparently un-mysterious aspects of bodily experience – the conscious experience of resistance when the child pushes against a stationary object or our experience of movement through space as we walk unimpeded – and attempts to show that all of our conscious experience is constructed out of these basic bodily materials.³

The point of much of this work seems to be to deflate the idea of abstract reasoning, abstract concepts, or amodal symbol systems (Glenberg 1997, Barsalou 1999, Lakoff and Johnson 1999), and this goal may seem to pertain more to matters cognitive (inference, concept-acquisition, online problem-solving) than to consciousness studies. But, notice the parity between the strategies employed in cognition-related domains and those pursued in connection with consciousness; Lakoff and Johnson explicitly take the sensorimotor domain to provide the raw materials both for inference and for conscious experience in other domains (understood via metaphorical extension from sensorimotor

³ Sometimes it is claimed that conscious, body-related experience solves the grounding problem (Harnad 1990, cf. Searle 1980) which might represent a version of this strategy (Lakoff and Johnson sometimes use the language of 'grounding', e.g., at 1999, 63). Others who take all of thought and experience to be somehow sensori-motor based (e.g., Barsalou 1999, Prinz 2002, 2012) might also fall into this camp.

domains). This cognition-related project is an attempt to show that the content or processing profile of even the most abstract concept – its role in inference, for instance – consists in nothing more than bodily materials re-used, re-enacted, or re-deployed. Thus, the success of this cognitive project entails successful appeal to a basic set of materials that can be combined to compose all of the target cognitive phenomena. Concepts of syntactic categories may ultimately be nothing more than, say, attractors in a dynamical system, attractors that formed in response to bodily experience (Petitot 1995, Rupert 1998). Mathematical concepts might be nothing more than motor routines deployed to control our discourse about mathematics (Lakoff and Nunez 2000). Notice the strength of the claim on the table, that *all* of cognition can be explained in sensorimotor terms or in terms of bodily activity. Some discussions seem tame by comparison, contemplating only the possibility that embodied experience constrains, or even more weakly, happens to facilitate, the acquisition of concepts and inferential structures. Lakoff and Johnson, for instance, sometimes use the language of dependence, arising from, being tied to, being based in, and being connected to (1999, 77; cf. Shaun Gallagher's use of such metaphysically noncommittal phrases as 'tied to' and 'comes along with' when talking about perceptual awareness and its relation to embodied processing – 2005, 74), but such weak claims are beside the point in the present context. If one hopes to show that embodied reasoning is the only kind, then no residue can remain. If one thinks that what appears to be the human use of abstract concepts is nothing more than metaphor-based appropriation of bodily materials, then one is committed to a reduction of sorts; all of our abstract concepts are reduced (by the process of metaphorical extension) to basic bodily materials.

This point about reduction applies even more straightforwardly to consciousness and the hard problem. Chalmers accepts that consciousness, in some sense, arises from or is connected to physical matter (e.g., the brain); the question is “why should it?” One thing’s merely arising from another leaves open the possibility that former has properties beyond what it arose from. In the case of consciousness, then, when conscious experience arises from physical matter, Chalmers wants to know what that extra bit is and how the arising of that extra bit is to be explained, given that what it arose from – the physical matter – is (apparently) of such a different nature. As much as one might want to avoid the language of reduction, then, a reductive project is the only option here. Thus, I’ll call the set of bodily materials a ‘reduction-base’, asking the reader to disregard what might be considered the unsavory history of the concept.

It may now be clear how the embodiment theorist might pursue Strategy 1. in an effort to solve the hard problem, at least on a case-by-case basis. If one wonders what constitutes ‘what it’s like’ to experience the linear passage of time, one can hold that such experience is entirely constituted from elements of the reduction base, for instance, bodily experiences of linear movement through space (Lakoff and Johnson 1999, 141). Or, if one wonders how it is that we can visually experience the distal nature of our physical world, one might claim that “distal experience is inscribed in bodily experience” (Gapenne 2010, 209) or that “the distally perceived...object is nothing else than the experience of body motion” (Stewart, Gapenne, and Di Paolo 2010, ix).

2. Subconscious bodily processes are basic; all conscious experience is shaped or determined by this basic stock of bodily processes.

On this view, bodily processes structure or precondition what subjects experience consciously; to borrow a colorful phrase from Lakoff and Johnson, bodily processes play the role of a “hidden hand that shapes conscious thought” (1999, 12). This strategy typically amounts to a kind of explanatory strategy. We have conscious experiences we would describe as have certain kinds of structure or texture – in the way in which they change over time, in the various peaks and troughs of intensity, in their shades of consciously experienced meaning. The structure of the subconscious processes that somehow cause, produce, or influence (that is, precondition) those conscious experiences explain the structure and texture of the conscious experiences.

The second and third passages quoted above from Gibbs seem to manifest this strategy; the temporal profile and other structural aspects of subconscious processes and interaction with the world cause our conscious experiences to be similarly structured, the former structure thus explaining the latter.

Some of Shaun Gallagher’s claims about the contribution of the body schema to consciousness seem to reflect strategy 2.; although not explicitly focused on the hard problem, Gallagher intends to explain certain facts about our conscious experience by appeal to subconscious body-related processes that structure or shape conscious experience. According to Gallagher, the body schema “is a system of sensory-motor functions that operate below the level of self-referential intentionality. It involves a set of tacit performances – preconscious, subpersonal processes that play a dynamic role in governing posture and movement,” the activity of which can operate “completely outside my awareness” (2005, 26). Even in cases in which I am attending to my body, “perceptual awareness of my body will always find its complement in capacities that are

defined by the operations of a body schema that continues to function to maintain balance and enable movement. Such operations are always in excess of that of which I can be aware” (*ibid.*, 27). The body schema “involves a *prenoetic* performance of the body. A prenoetic performance is one that helps to structure consciousness, but does not explicitly show itself in the contents of consciousness” (*ibid.*, 32; and see 38, 73). And, in the context of a discussion of early development and the neonate’s tendency to imitate facial expressions, Gallagher says, “Since the visual perception and proprioceptive awareness involved in such imitation are in fact instances of conscious experience, then consciousness is from the very beginning structured by embodiment in certain ways” (*ibid.*, 78, and see similar comments at 106).

At some points, too, Gallagher seems to rely on a premise that serves as a presupposition of strategy 2., that the character of conscious experience should be explained by subconscious, subpersonal processes. Consider, for example, his extended critical discussion of the theory-theory and the simulation theory of our understanding of others’ (and in the case of the theory-theory our own) minds. Although Gallagher acknowledges (*ibid.*, 209, 212) that the advocates of these approaches hold the cognitive processing in question to be subconscious, Gallagher nevertheless grounds substantial criticisms in conscious experience: it doesn’t seem to us as if we’re explaining and predicting things when we interact with others, and to Gallagher’s mind, that counts substantially against the theory-theory and the simulation theory (*ibid.*, 212). This reasoning makes little sense unless Gallagher assumes that conscious processes reflect subconscious ones, and that a theory of subconscious processes that produce a certain kind of behavior is worse off than one that matches conscious experience or has its

structure reflected in it – presumably because the former doesn't explain conscious experience as well as the latter. And, returning to the hard problem, this reasoning itself seems to presuppose that a substantial explanatory relation can hold between subconscious processes and conscious ones (cf. Noë 2004, 225–226).

Michael Wheeler makes a similar move in his *Reconstructing the Cognitive World*, treating “structural isomorphism” between the conscious and subconscious levels as an evidentiary lever (2005, 227). His central argument for extended cognition runs as follows: dynamical-systems-based descriptions of the subpersonal level extend into the environment in some cases, and they do so in just those cases in which we have no conscious experience of a subject-object boundary; thus, the extended dynamical-systems-based view of these cases is phenomenologically in key (more so than any nonextended view). Thus, the extended dynamical-systems-based model is more likely to be correct than a rules-and-representations-based account of the cognitive processes at work in such cases. And, so cognition is sometimes extended (*ibid.*, 226–227). Here Wheeler presupposes an explanatory relation of just the sort Gallagher seems to assume, at least with regard to the content of conscious experience, for, as Wheeler would have things, the DS-based view wins by inference to the best explanation; the relevant conscious experience is better explained by DS-based view than it is by the (nonextended) computationalist one. Since both of those views are theories of what happens at the subconscious level, Wheeler would seem to assume that the subconscious processing explains the nature of conscious experience (although the hard problem of conscious experience is nowhere mentioned in Wheeler's book).

3. *What might normally be thought of as subconscious bodily experience constitutes or is identical to conscious experience.*

In its most straightforward incarnation, Strategy 3. rewrites the explanatory stories of Strategy 2. as odes to constitution. One's conscious experience of planning to build a chest of drawers isn't merely shaped or preconditioned by the activation of motor routines for swinging hammers and the like, routines we would normally take to be subconscious; such activated motor routines literally are the conscious experience of planning. It's as if we had (a) lined up all of the merely mechanical, subpersonal causal contributors to the production of the sorts of behavior of interest to cognitive scientists and philosophers of mind – none of which contributors would, viewed in this light, be treated as a conscious experience – and then (b) specified conditions that, when met, confer on some of the states or processes on the list status as conscious experiences.

I extract this view partly from Chemero's work, as quoted above, and I think it's to be found in writings of other embodiment-oriented theorists with enactivist leanings, that is, who are inclined to think that conscious experience just is some kind of skilled bodily action or interaction with the environment. Gallagher sometimes seems to place himself in this company, at least with regard to a certain kind of conscious awareness: "The infant already knows how to grasp, and is performatively aware of its hand *in* the grasping" (2006, 74).

Or, consider the following claim, made by Alva Noë:

We don't *apply* sensorimotor knowledge *to* experience. Rather, we bring it to bear *in* experience; bringing it to bear in this way enables what would otherwise be mere sensory stimulation without world-presenting content *to be* experience. Perceptual experience *just is* a mode of skillful exploration of the world. (2004, 194)

On one reading of this view, this illustrates Strategy 3. perfectly. Our perceptual experience – the paradigmatic sort of experience that, according to Chalmers, generates the hard problem – is a matter of exhibiting certain forms of behavior or interacting with the environment in a certain way. Perceptual experience just is the exercise of knowledge of sensorimotor contingencies, and that knowledge just is the failure to engage in “surprise” behavior during the subject’s sensory interaction with the world, and thus, the pattern of behavior itself (and perhaps the neurophysiology at work) is the conscious experience.

Some contributors to a relatively recent volume of essays on enactivism pursue a similar strategy. How can the mind-body problem be solved? According to John Stewart, “The paradigm of enaction solves this problem by grounding all cognition as an essential feature of living organisms” (Stewart 2010, 1). How could this be? In a later essay, readers are told that “[i]n sense-making, active coupling with the world brings forth a realm of significance” (Di Paolo, Rohde, and De Jaegher 2010, 71). Now, given the normal acceptance of these terms, their readers will likely take the sense and significance in question to inhabit a realm of abstract ideas or an internal realm of mental states (perhaps conscious experiences!) and take the relation between such states and human action to be an open and important question. But, for enactivists of the sort in question, ‘sense-making’, ‘significance’, and ‘meaning’ are all defined as certain kinds of interaction between the organism and the world, in ways that further life (*ibid.*, 36, 39, Colombetti 2010, 150). To say that our cognitive interactions with the world create significance, then, is to say that our activities of interacting with the world in ways that further life are activities of interacting with the world in ways that further life!

In the end, the view of consciousness on offer would seem to be a straightforward constitution or identity claim; behavior and what we would normally think of as the subconscious processes that produce behavior, together with the way that the world pushes back to alter the internal processes in question, constitute or *just are* conscious experience.

Bear in mind that my project in this section has not been exegetical. Many embodiment-oriented authors don't make it sufficiently clear which way they mean to attack the hard problem or whether they intend their accounts of consciousness to solve the hard problem at all. I hope only to have convinced the reader the embodiment-based literature contains at least these three ways of thinking about consciousness, and I also ask the reader to trust that I found no others.

Section 2. The Hard Problem Strikes (the Body) Back

In this section, I argue that all three embodiment-based strategies identified in Section 1 seem hopeless as head-on solutions to the hard problem. I work through them in order.

1. *Conscious bodily experience is basic; the remainder of conscious experience is constructed from this basic stock.*

If one hopes to explain the appearance of mysterious or impressive-but-not-understood phenomenon X in the universe, then it is beside the point to propose that a certain amount of X can engender more. How do we understand love? As a journey. What constitutes our understanding of journeys? Our experience of bodily motion through space. What gives *that* experience its “what it’s like”? Deafening silence. The hard problem remains

unsolved, for what it's like to have bodily experience is left unaccounted for.⁴ Zombies don't have it, Chalmers would insist, yet they have all of the physical and functional structure of a smitten human.

To be fair, discussions in this vicinity reward careful attention. Lakoff and Johnson's appeal to metaphorical extension may well explain why my experience of love has some of the particular aspects it has. Perhaps my situation now of having gotten off track in my relationship with my romantic partner feels the way it does because I've built my romance-related conscious experiences – both positive and negative – from my conscious experiences of journeys; I was once lost on a trail while hiking with my parents as a child, and perhaps this contributes. This might constitute a revelation of sorts to me. Nevertheless, it's no answer to the hard problem. Why did getting lost on the hike with my parents feel *that way* instead of feeling the way it feels to win the lottery or eat an orange or hear a symphony? Nothing about the physical process of my having gotten lost or the structure of the causal processes that occurred at the time explains why being lost would feel one way rather than another. Why wouldn't those physical or structured causal processes feel like winning the lottery rather than being lost? The essential problem is that the reduction base alone causes the hard problem. So, reducing everything else to the base does nothing to remove the mystery of conscious experience, although it may help us to see more clearly what exactly those who take the hard problem seriously should focus on.

⁴ Lakoff and Johnson offer a more subtle picture of love, making room for some basic experience of love (1999, 70). But, this nuance does nothing to obtund the force of the criticism made of Strategy 1.; regarding the part of our understanding that consists of a basic experience of love, what Lakoff and Johnson offer is a restatement of the hard problem – love feels like *that* – but no solution.

That being said, the reductive step itself threatens to be a problem; the full derivation of nonbasic conscious experience from basic bodily experience is no mean feat. Face it. Attempts at this sort of reduction have a dismal track record. There's residue even in the best cases (think of the challenges faced by Russell's phenomenalism or Hume's classic empiricism – Rupert 2011, chapter 11). Consider the cognitive analogy: If I want to get my romantic relationship back on track, I don't typically consult a Google Maps application, which I would consult if my actual journey were off track. The point is that for any target domain for metaphor, the domain seems likely to have at least some of its own distinctive inferential or experiential content. In which case, poof goes the reduction. Of course, one can try for a complex story, piecing together the different cognitive and experiential aspects of each domain from basic bodily experience. But, this is no easy road, and one shouldn't assume that, simply because Lakoff and Johnson offer many examples of partial overlap, complete overlap or nesting awaits discovery.

2. 2. Subconscious bodily processes are basic; all conscious experience is shaped or determined by this basic stock of bodily processes.

I have no doubt that subconscious bodily experience causally affects what we take to be conscious experience – or at least affects the behavioral output we take to be caused by conscious experience – but this observation doesn't begin to address the hard problem, as posed. Instead, it serves only to make the hard problem more pointed. Why, we naturally ask, should *that* bit of subconscious processing give rise to *this* sort of experience – not just experience that's structured like this, but conscious experience with this particular phenomenological feel. It's one thing to say that the temporal or spatial structure of

experience arises mirrors aspects of the temporal or spatial structure of subconscious bodily processing. It's another thing to explain why experience with a particular temporal structure feels like this rather than that (why the experience of the passing of time during, say, motor preparation, feels like the passing of time during motor preparation rather than feeling like the passing of time in free fall from an airplane). Structure is of great interest, but it must be imposed on something. The hard problem is that of explaining the nature of the things on which structure is imposed, not the structure itself.

Setting aside concern about the hard problem, one might transform Strategy 2. into a constraint on the selection of models of subconscious- or subpersonal-level processing. If underlying structure gives rise to conscious structure, then data about the latter constrains the selection of theories of the former, doesn't it? Not so quick. It may seem natural that we would find, in whatever processes give rise to conscious experience, the same structure we find in that experience, but conscious experiences could just as well be produced by subconscious processes that aren't structurally similar to conscious experiences as by ones that are; it depends on the nature of the processes by which subconscious states give rise to conscious ones. Wax imprints bear structural marks of their causes, and we can see why once we know something about the imprinting process; but explosions don't. Likewise, we should be open to the possibility that a model of subconscious processing that explains the production of first-person verbal reports does not invoke states with the same structure as (or that are accurately described by) those verbal reports. Which kinds of models we endorse, matching or mismatching, depends on how well those models fit the data, the extent to which the models dovetail with useful

models of related processes, and so on, through the standard desiderata for successful modeling and theory choice.

3. What might normally be thought of as subconscious bodily experience constitutes or is identical to conscious experience.

As a way to solve the hard problem, this strategy is both a nonstarter and a path forward. On the one hand, by invoking what is normally thought to be subconscious processing (or overt behavior, and thus, one might think, neither conscious nor subconscious), the proponent of Strategy 3. seems only to widen the explanatory gap. How could that stuff – interaction with the environment, for example – *be* phenomenal experience? It doesn't *seem* at all like consciousness. Moreover, the internal processes described are, on their face, merely mechanistic processes of the sort that anyone who takes on the hard problem, as posed, has agreed to treat as beside the explanatory point.⁵ On the other hand, an important insight seems to animate the proponents of Strategy 3.: that consciousness isn't what we thought it was. When properly developed, this supposition grounds not a scientific or materialist solution to the hard problem, as posed, but a materialist dissolution of it. Let me explain.

Authors who pursue Strategy 3. often focus on the nature of life and biological needs or the organism's fittedness to its environment. In the former case, the idea seems to be that consciousness is, by its nature, pointed toward an object, and such object-pointed processes – intentional processes, as they're often called – provide consciousness, or at

⁵ See, for example, Noë's explanation of how his enactive view avoids circularity – by appealing only to facts only at the neural and information-processing levels, so as to maintain scientific credibility – and why it is thereby inadequate: "But how," he asks, "can phenomenally unconscious states of this sort be the basis of phenomenal consciousness?" (2004, 228–229)

least proto-consciousness, from the time they appear. In the latter case, the organism's fitting interaction with its environment – either the organism's skillful or knowledgeable interaction with the environment or the organism's making use of Gibsonian affordances to guide its action – constitutes consciousness.

And some authors entertain both thoughts. Noë sometimes seems to place himself into the latter camp, offering a “skilled interaction” account of conscious perceptual experience: “Experience is realized in the active life of the skillful animal” (2004, 227). But, when confronted head on with the mystery of consciousness and its appearance in nature, Noë goes biological (*ibid.*, 229–230), arguing that “with increasing sensorimotor complexity you get the appearance of a life form that embodies a measure of sensitivity to the way its own movements change the way the environment stimulates it” (*ibid.*, 230).

These seem hopeless to me as head-on responses to the hard problem. The maintenance of the organism in the face of a shifting environment – by metabolic processes, cell division, and all the rest – is an incredibly complex affair, as is the way in which the organism navigates its environment so as to keep these processes running relatively smoothly. And, there is certainly something “directed” about such behavior, whether thought of in real-time or in terms of the evolutionary history of the mechanisms that control such processes. But, this the wrong sort of complexity from Chalmers's standpoint. It's a complex network of causal and functional facts. These are the wrong sorts of facts to explain the intrinsic qualitative character of conscious experience and thus adding more and more of them, and more patterns among them, simply does no good. The problem is that those sorts of facts – facts about causal, functional, and dynamic processes – are relational. They concern properties or characteristics identified

in terms of their relation to other properties or characteristics and thus can't explain why our experiences have the intrinsic qualities they do; those qualities are nonrelational.

Perhaps some kind of gradualism will help (as Noë's comments seems to imply). At the lowest degrees of complexity (or earliest stage in the development of the organism), there is no consciousness. But, as complexity increases, consciousness emerges.

Questions about complexity, emergence, and self-organization are fascinating, but I don't see how an exploration of these issues can promise any progress on the hard problem; for the concerns of the previous paragraph apply equally to this gradualism. There may be a gradual shift toward greater complexity, one that yields new materials, but only more materials of the same unhelpful kind. Perhaps gradual shifts eventuate in a sudden shift of sorts, allowing for new function: perhaps neural or other physical structures can suddenly become complex enough guide behavior in the absence of the environmental items that those structures normally track in real time. Alas, though, we are given only more function; only a relational matter has changed, and we have nothing with which to explain the intrinsic, nonrelational features of conscious experience.

The gist of the present section is that the hard problem is hard, in fact, too hard to be solved head-on by appeal to embodiment. That being said, the instincts behind Strategy 3. point us in the right direction. They encourage use to shift our attention to the scientific modeling of the data we associate with consciousness or that seems relevant to our attributions of conscious states – to open our minds to whatever might account for the phenomena of interest and to the possibilities both that the phenomena and whatever accounts for them might not have the characteristics we expected.

Section 3. Neurophenomenology

Some advocates for the embodied view propose a different response to the hard problem, not a head-on solution, but instead a “remedy” (as Varela [1996] calls it in the title of his agenda-setting paper on the topic). On this view, embodied cognitive science will solve problems to do with consciousness by taking seriously subjects’ first-person reports concerning their own conscious experience. In particular, this method recommends that cognitive scientists train subjects in phenomenological reflection, which allows them to report accurately, not so much on the contents of conscious experience (e.g., “I seem to see a table”), but on something more elusive, “our capacity for being conscious” (*ibid.*, 335); this sort of reflection reveals the structures that constitute a kind of pre-understanding that precedes, shapes, and suffuses experience (*ibid.*, 336). This approach adds an essentially first-person perspective – the perspective of lived experience – to cognitive science’s standard third-person approach.⁶ To be clear, Varela does not recommend that we abandon standard cognitive-scientific tool – fMRI magnets, eye-trackers, computer software that records reaction times in behavioral experiments, and the like. Rather, he plumps for a new kind of cognitive science, one that weds the first- and with the third-person perspective reflected by the use of these standard tools; and this is an embodied cognitive science precisely because our bodily experience creates or forms the pre-reflective structures that shape and suffuse first-person conscious experiences (cf. Gallagher’s position, discussed above in connection with Strategy 2.).

As Varela conceives of this neurophenomenological method, the first-person perspective adds no mere bell- or whistle-stream of additional data. It adds irreducibly

⁶ It should be noted that Varela is not entirely happy with this language, given that the so-called first-person view is suffused with social dimensions (and is thus third-person, in a sense) and that scientific method is itself up to its ears in social dimensions that don’t fit neatly with the kind of objectivity or neutrality that many have in mind when they talk about the third-person perspective (Varela 1996, 340).

first-person *explananda* and a uniquely first-personal flavor to cognitive science.

Characteristics of the phenomenological ground are aspects of the very phenomena to be accounted for by cognitive science, and they are directly reflected in, say, first-person verbal reports on the prereflective structures that shape consciousness.

What does neurophenomenology look like in scientific practice? In what follows, I'll describe a parade result (Lutz, Lachaux, Martinerie, and Varela 2002), which provides a case study, to be used as a neurophenomenological reference point throughout the remainder of this section and the next. In this study, subjects view a computer screen with a series of two random-dot stereogram, first one that cannot be fused, followed by one capable of being fused. An auditory warning sounds when the second stereogram (with a binocular disparity in the pattern that allows for fusing and an image to “pop out”) appears in place of the first. Subjects are instructed to press a button as soon as they see the hidden shape appear. In pre-experiment, training trials, subjects are asked to reflect on their own mental states and processes prior to and during the trials. Experimenters repeated this procedure until all four subjects had arrived at stable categories into which to place new trials. From these categories, the experimenters constructed phenomenological clusters (PhCs) of trials, including, most importantly, Steady Readiness and Unreadiness. For the experimental trials, subjects were fitted with standard electrode caps, and EEG recordings were made. Subjects were run through hundreds of the same kind of stereogram trials as described above and were asked to categorize them, which allowed the experimenters to separate experimental trials into PhCs.

The important findings were (a) that differences in subjects' self-reported frames of mind – and the corresponding PhC categories – accounted for a significant amount of

variance in their reaction times in response to the appearance of the fused figure and (b) that subjects' categories were also statistically correlated with various aspects of the EEG data, including long-range and short-range synchronization of firings recorded at individual electrode sites, as well as the distinctive EEG patterns evoked by the stimuli.

What do these results tell us about neurophenomenological inquiry? We seem to have a neurophenomenological success, but success in what respect precisely? According to the study's authors, the study extracts "data from subjective experiences," paying "meticulous attention to the intimate and direct knowledge that a subject has about his/her experience" that "can effectively constrain the analysis and interpretation of neurodynamical data" (Lutz et al. 2002, 1590). These are high-flying claims. Are alternative glosses available? Are they more compelling or more justifiable than the authors'?

Varela's vision and the experimental execution of it suggest at least the four following ways to conceive of neurophenomenological research:

1. *First-Person Explananda*. Unreduced conscious experiences are explananda in cognitive science. They are on par with such data as reaction times collected in behavioral experiments (they have "equal status" – Varela 1996, 343), not part of the theoretical machinery (e.g., computational models, connectionist networks, neurobiological models) introduced to account for such data as recorded reaction times.

2. *First-Person Access*. Access to those explananda comes through subjects' trained reports, which are assumed to be veridical in their content; if a phenomenologically

trained subject says, “My conscious experience has structuring properties P and Q,” then the subject has (or is very likely to have) conscious experiences with structuring properties P and Q, and the subject’s being in that state becomes a data point for cognitive science (in contrast to that state’s being treated only as part of a hypothesized model meant to explain further data, such as verbal reports).

3. *First-Person Perspective*. This expresses a watered-down vision of *First-Person Access*. On this view, there is something specially first-personal about phenomenological reports collected in neurophenomenological experiments, but *First-Person Perspective* makes no metaphysical or ontological claim. Rather, the point is that phenomenological data introduce a genuinely new, first-person method (one that reflects “lived experience” – Varela 1996, 345) into cognitive science and, most importantly, the distinctive first-personal status gives the first-person data a kind of trumping power or, if not quite that, a privileged status which entails an extra-forceful constraining power: first-person data must be protected and respected in way that standard scientific method does not underwrite; first-person data has to be given more weight, in the process of the modeling of data, than would normally be given simply to a stream of collected data, such as a set of reaction times (Varela 1996, 344, Lutz 2002, 135).

4. *Methodological Tweak*. On this view, neurophenomenology consists in nothing more than the design of new kinds of experiments, which may happen to be inspired by, say, the writings of Husserl or Merleau-Ponty. This approach does not merely avoid the hard problem or in some other way dance around it (as the approaches others do); it

unabashedly sets to work on only the so-called easy problems of consciousness. This conception of neurophenomenology casts it as normal cognitive science. It recommends training subjects to make certain kinds of reports, but this training does not differ in any essential methodological regard from standard pretrial training that takes place in all manner of psychological experiment (to be sure the subjects understand the instructions, for example, or can use the experimental devices in question). On this view, the first-person reports produced are subject to standard statistical analysis; that is, they are treated as simply one more stream of data – along with neuroimaging and other behavioral measures – the modeling of which might reveal cognitive architecture, processes, or mechanisms. The first-person reports have no special status and their content is not presumed (prior to the results of statistical analysis and successful model building) to describe accurately internal processes of interest.

Some further comments are in order concerning the entries on the list. Keeping our eyes on the hard problem, we might note that, in combination, entries 1. and 2. differ little from Chalmers's own way forward in consciousness studies.⁷ Varela recommends that, in response to the hard problem, we search for “meaningful bridges between two irreducible phenomenal domains” (1996, 340); presumably he has in mind the domain of the sort of conscious observation that drives scientific enquiry, on the one hand, and the domain of trained phenomenological report, on the other. Similarly Chalmers holds that facts about the what it's like of conscious experience cannot be derived from the sort of observational facts associated with empirical enquiry, and this reductive failure itself

⁷ Here I bracket Chalmers's interest in Russellian monism and his suggestion that a single information space limns both physical and phenomenal realities.

helps to create the hard problem and his dualist conclusions (Chalmers 1996). What method of enquiry does he recommend in response? Chalmers holds that contingent psychophysical laws hold in our world, that they ground various structural isomorphisms, and that we can investigate those laws by investigating conscious experience in tandem with our investigation of neural and computational structure and function (Chalmers 1995). That being said, Varela seems less concerned than Chalmers with ontology and more concerned with what he calls pragmatic issues – prioritizing methodology over metaphysics.

Second, I'll remark on the nature of a neurophenomenological science. Roughly speaking, the entries on the list are placed in descending order of their “mysteriousness,” from the standpoint of the orthodox cognitive science. Perhaps most obvious is the divide between *Methodological Tweak* and the other three, with regard to their revolutionary credentials; *Methodological Tweak* has essentially none, while the other three encourage cognitive science to become a deeply different sort of enterprise, one that, to many minds, might not seem like a science at all (normally no data stream has independent trumping power and no nonpublic data count as *explananda*, except derivatively). How should neurophenomenology's apparently anti-scientific outlook be understood?

The neurophenomenologist seems to find solace in two theses. First, Varela claims that the hard problem simply cannot be given an “intellectually coherent” solution (1996, 345) without genuine incorporation of our lived experience into the development of our account of consciousness; and in his view, lived experience can only enter into the project in a meaningful way when it's given something like the independent, or even trumping, power recommended in various ways by *First-Person Explananda*, *First-*

Person Access, and *First-Person Perspective*: “We are similarly convinced by empirical and intuitive evidence that our human experience, mine as well as yours, follow [sic] some fundamental structural principle which, like space, enforces the nature of what is given to us as contents of experience” (*ibid.*, 340)

This quotation may convey some of the reason why and how, according to Varela, our lived human experience must be respected by scientific investigations, and also why he tends not to think of this as a commitment to dualism. For, the quotation introduces the second thought, which concerns the self-reflective nature of neurophenomenology. According to this view (cf. Borrett, Kelly, and Kwan 2000), science itself must arise out of the structure of consciousness that is directly explored in phenomenological reflection. Thus, phenomenological reflection is, as much as anything, an examination of the very foundation of science, in connection with such fundamental categories of conscious experience as those to do with time and space (which are themselves conditioned by bodily processes and experiences). This vision accounts for many of the views described above in connection with neurophenomenology. For instance, it accounts for the privileged role that many neurophenomenologists assign to trained first-person reports; these are insights into the very materials out of which science grows and in terms of which it must be framed. Thus, it is no wonder that such reports have a privileged status; it would be not only a mistake not to accord them such status, but an act of hubris or extreme naiveté. Return now to Varela’s point about intellectual coherence: if the structures revealed by phenomenological reduction are themselves the roots of scientific thinking and observation, then, when science turns to the study of consciousness itself, it must produce a story that jibes with the output of phenomenological reduction; not only

should the two jibe, but the scientific story must confer a central role on the phenomenological experience itself, that is, science must recognize, in its own theories, the foundational role, in the construction of science itself, of phenomenologically revealed characteristics of conscious experience. Any account of consciousness that doesn't thusly relate its own output to the output of phenomenological reduction is, in Varela's words, intellectually incoherent. This solution is deeper and more satisfying, in Varela's view, than the sort of dualist naturalism proposed by Chalmers.

One should wonder, though, what to think if, upon doing the most rigorous and thorough nonphenomenologically based cognitive science, our best theories tell us that, and why, our reflective experience is actually misguided? What happens if the science that we thought was a product of conscious observational experiences structured in a certain way (which we think was revealed to us by phenomenological reflection) tells us that those phenomenological reflections are, say, misleading by-products of the processes that actually allow us to construct successful scientific theories? What if we do a cognitive science of the production of phenomenological reductions and it tells us that the outputs of such reflections are erroneous and explains why they are? These questions set the stage for the critical discussion of Section 4.

Section 4. Neurophenomenology: Neither Science nor Phenomenology?

In this section, I argue that neurophenomenology is either an unmotivated drastic departure from standard scientific assumptions or a low-profile and somewhat humdrum (although still interesting) affair along the lines of *Methodological Tweak*.

Consider the first three strands of neurophenomenological thought described above: *First-Person Explananda*, *First-Person Access*, and *First-Person Perspective*. I am

puzzled by the insistence that such views have a place in the scientific enterprise, particularly given that (a) they seem to presuppose dualism (even *First-Person Perspective*, for it posits a methodological exceptionalism without any rationale, and what might that be other than a dualist one?) and (b) *Methodological Tweak* accommodates perfectly well such experimental successes as the ones had by Lutz et al. (more on which in connection with the second horn). In the sciences, *explananda* are either fine-grained – specific data sets – or what we take to be coarse-grained phenomena – large-scale patterns in our observations (e.g., there seems to be a temperature cycle on Earth that runs cooler for a while, then hotter, then cooler again). But, regardless of whether the explananda are fine-grained or coarse-grained, they are measurable and publicly observable. So, although many people trained in a phenomenological method might say that they have experiences with such-and-such structural properties, the data is the saying so, not the experiences themselves. Cognitive science should stick to that methodological commitment, unless very strong pushes elicit very strong shoves. The strongest support for this “scientistic” position rests on a relatively mundane observation: throughout the course of human history, science has marched on precisely by ignoring (and sometimes having to fight actively against) the voices of those who want to introduce or retain all manner of pretheoretical commitments to the supernatural, mystical, un-measurable, and so on.

What of the transcendental view, though, according to which it is only the prereflective structures available to the trained phenomenologist that make science possible? The straightforward response is that the scientific community isn’t beholden to this philosophical vision. Consider the following comparison: Joe takes his relationship

with a Christian God to structure all of his experiences. Thus, Joe takes seriously a model of cognition only if that model incorporates the contributions of his relationship with a Christian God to the production of his conscious awareness; perhaps Joe's conception of a Christian God puts that God both inside and outside time, and thus the only models of cognition legitimately on the table for Joe are those that respect his reflective sense that he walks with God both inside of time and outside of time. Surely, Joe's restrictions on legitimate contending theories of cognition and consciousness have no force in cognitive science. To claim otherwise is a nonstarter. I assume such a reputable scientific journal as *PNAS* published the paper by Lutz et al. precisely because the methodology of the experiment itself made no commitment to any such restrictions, even if a bit of the commentary (where much more latitude is allowed in the sciences) suggests a commitment to them.

To be clear, of course cognitive science must ultimately explain how it is that cognitive science gets done by human minds; it must be self-reflective in the sense that it must be self-subsuming. For example, a computational theory of vision had better explain how it is that human visual processing allowed experimenters to track the data (reading the instruments, etc.) to which that theory was epistemically responsive. This, however, leaves open many possible options for cognitive-scientific modelers; moreover, a properly scientific investigation will not allow the exploration of such various models to be truncated or constrained by the demand that a particular philosophical program be vindicated.

How, then, should we understand what's happening in the experiments performed by Lutz et al.? Might they deliver something genuinely scientific yet also uniquely and

substantively neurophenomenological? Begin with some of the central elements of what Gualtiero Piccinini (2010) calls ‘self-measurement methodology’. Piccinini claims that first-person reports are evidence of something, but not necessarily of anything specially to do with consciousness; first-person reports are data points that have the potential to indicate something about the subject, but whether they’re *about* conscious states “is best left to a theory of consciousness” (*ibid.*, 88) to decide. Moreover, Piccinini claims, our category of interest in these contexts should be the broad category of first-person behavior, rather than its species first-person report: “Thus, ‘first-person report’ is too restrictive to capture all sources of first-person data. Instead, I use ‘first-person behaviour’ to denote any behaviour that is a source of first-person data” (*ibid.*, 89).

The second point is important, because any attempt to stop short of *Methodological Tweak* is likely to emphasize the first-personal nature of the reports. Such stopping short might consist in an attempt to formulate a *MT+*, a view that, for the most part, embraces standard scientific methodology but adds there’s something genuinely innovative – more than a tweak! – involved in the first-person method. Now, as we’ve seen, Piccinini emphasizes the sense in which first-person behavior (e.g., button-pushing) is data that, like self-report, issues from the subject and might indicate something interesting about the subject’s cognitive processing, consciousness, or mental states. I think some authors have taken first-person report to be special partly because such reports have content.⁸ But, to bolster Puccini’s picture, I would argue that our putting reports on par with other kinds of behavior requires that we all such behavior, including the reports, as something like causal signals of some internal states (the ones that produced the output!), and this is

⁸ Although, against the backdrop of certain instructions given by the experimenter, an action such as a button-pressing might report, that is, it might carry content.

difficult to do given a striking and distinctive aspect of first-person reports: that they have content. It's worth flagging, then, the many and varied reasons we have to be skeptical about that content as conveying an accurate picture of the processes that produced it. The evidence in this regard is wide-ranging and copious in quantity, from the wealth of results surveyed in Nisbett and Wilson (1977) to the wide range of more recent results filling Carruthers (2011) (also see Kahneman 2011, Gertler 2010, chapter 3, as well as Puccini's compelling examples taken from the work of Johansson and collaborators – 2010, 97–98). Of course, neurophenomenologists can insist there's something special about the first-person reports of trained phenomenologists, with regard to the structural of conscious experience. But, given the wealth of data suggesting the unreliability of introspection, this would seem like an article of faith, a promise to “trust us.”

So, how might we think of first-person behaviors? According to Piccinini, they are causal indicators of internal states. This is nearly a truism. Any form of behavior carries information about something, typically the process that produced the behavior. Movement of the arm carries the information that the end plate was stimulated, for example. Similarly, any verbal report exhibits various qualities: some of which could be mere physical properties, such as loudness etc., which reflect the processes that produced it. Whether the publicly understood content of the report correlates with anything more than a motor command to produce those sounds remains to be investigated. It can be proposed as a tentative hypothesis that a self-report has content that accurately describes the processes occurring in the subject's cognitive system, but it would be subject to normal testing and verification (and nothing in Lutz et al. results suggest otherwise).

That being noted, Lutz et al. provide more than a trivial reflection of a cause by its relatively immediate effect. The studies at issue are of interest because there is, so to speak, a four-way connection: First-person reports correlate with certain processes that casually contribute to them, and those latter states are correlated with the stimulus-evoked neural changes, which are themselves correlated with subjects' behavioral reactions. Thus, we're given evidence that the processes that produce reports modulate processes that causally contribute to the production of the behavioral report (the subjects' pressing of the button when they see the geometrical figure appear in the stereogram). But, notice the extent to which *First-Person Explananda*, *First-Person Access*, and *First-Person Perspective* all represent optional theoretical glosses of this result, given by the authors to experiments that have nothing much to do with those commitments. One needn't think there are any robustly conscious states to understand the methodology and results of Lutz et al. and to understand their force and find them interesting. One needn't think the subjects are directly reporting on conscious states. And, one needn't think there's anything especially first-personal about the data, or any more first-personal than the data one collects in any old experiment in which the behavior of the subjects is causally related to the input.

Thus, the success of these studies supports only *Methodological Tweak*. The kind of training provided to the subjects may seem unusual relative to the sort of pre-trial training normally given in behavioral experiments, and this is of genuine interest; perhaps more experimentalists will want to give it a try.⁹ But, this experimental work and the statistical

⁹ I suspect that Lutz et al. are impressed by the use of a mental state term during training, which suggests that they subjects were genuinely introspecting their mental states. One might reasonably wonder, though, whether the training corrupts the use of such terms or creates a specialized use for them. For my part, I'm not sure there is such a thing as attention, but be that as it may, we at least attempt to refer using 'attention',

analysis involved nowhere approximates the sort of constraint that many neurophenomenologists seem to have in mind; to be sure, the first-person data constrain the other data, but only in the conditional way that any data constrain any other data: to the extent that various factors in the data are statistically correlated, we should develop a model of the component processes that accounts for that correlation. This does not, however, given first-person reports any distinctive power to validate neuroscientific models of cognition or consciousness (1996, 344) – unless of course all Varela has in mind when he talks of phenomenological validation is the standard procedure just described (of performing statistical analysis on various data streams), in which case there is nothing to neurophenomenology beyond *Methodological Tweak*.

Section 5. Embodied consciousness and the self

Enough with all the naysaying. What, then, is the connection between embodiment-related results and our understanding of consciousness? Let me begin with stage-setting, as I see the stage (Rupert 2011). First, ‘consciousness’ is a natural kind term, as is ‘attention’,¹⁰ and thus we should begin our investigation of consciousness and attention

and we do so by employing various linguistic and pragmatic mechanisms, the primary ones of which involve untrained applications to ourselves as well as applications to other people engaged in what we think is attentive behavior. When asked to use the term in the laboratory context, during many trials of training, it might get contingently attached by subjects to some internal state of causal import which may not be the state everyday subjects refer to (if anything) when they use ‘attention’. (Compare: I’m trained in an experimental setting to use ‘dog’ to refer to some cartoon five-legged aliens. I may be able to perform quite well on the tasks I’m asked to perform – “tell me when the dog appears on the monitor” – even if ‘dog’ doesn’t refer to what it normally refers to or has its use governed by mechanisms other than those that normally govern its use.) And, what I say in this note about ‘attention’ would appear to apply to any mental-state term that might be used during phenomenological training.

¹⁰ Compare Prinz (2012, 90), who seems to think only the latter is a natural kind term, whereas the former, in contrast, is subject to conceptual analysis. And what is said conceptual analysis? On Prinz’s view, mental states are conscious if and only if “they feel like something or...there is something it is like to have them” (*ibid.*, 90). I would urge, though, that ‘feeling like’ is itself a natural kind term, which confers on ‘consciousness’ the relevant properties of a natural kind term (e.g., that one can’t know the essential properties of its referent *a priori*). At least, for those who make the bet, as Prinz does (*ibid.*, xi), that consciousness is part of the physical world and can be fruitfully investigated scientifically, I see no reason why we wouldn’t treat ‘feels like’ as a psychological term and thus a natural kind term – or, if not, as a

by probing the observable phenomena associated with the use of these terms (but not definitional of the terms). Second, most of what produces the behavior we associate with consciousness and attention is produced by a suite of cognitive processes (where ‘suite’ is a euphemism, used in place of ‘barely coordinated mess of thousands of processes running simultaneously, forming causal coalitions, competing, dissipating, and so on’ – Dennett 1991). Third, in virtually all cases, some distinctively bodily processes constitute part of the set of causal contributors to the production of the behavior associated with consciousness and attention. Here I endorse one version of a central aspect of the vision I attributed above to Varela, Gallagher, Lakoff and Johnson, and Borrett et al. – deriving from Husserl and Merleau-Ponty – according to which bodily processes shape our mental lives. Even when, for instance, the output of abstract reasoning processes initiate a motor command, various bodily processes also causally contribute to the issuing of that same command or causally affect the steps in the abstract reasoning process that led to the issuing of the command. Fourth, the kind of influence here reflects what I’ve elsewhere called a massively representational mind (Rupert 2011): our minds teem with representations, very many of which co-refer; they are distinct psychological vehicles, but they represent (in the externalist’s sense) the same kinds, individuals, or properties – in roughly the way that ‘cat’ and ‘gato’ do.

Let me now bring this more closely to bear on embodied cognitive science. Assume that, on various occasions, I interact more or less skillfully with the environment, whether it’s reading a book or doing the dishes. I contend that this is partly the result of having more or fewer content-sharing representations active at the same time (perhaps as a

dispensable term, at least in the context of cognitive science, in which case, (phenomenal) ‘consciousness’ would be likewise dispensable.

proportion of all representations contributing to the actions in question). The difference between the use of the dish towel to which a thousand different active neural units each of which represents the dish towel contributed, on the one hand, and the use the dish towel to which only ten representations of the dish towel contributed, on the other, accounts for some of the variance in the skillfulness of my use of the dish towel across these instances of use. Perhaps more surprisingly, I claim that such processes might take place internally. The process by which a subject arrives at a conclusion in reasoning may occur more or less skillfully or reliably, and that variation can be explained partly by the number of co-referring representations that contribute (as a proportion of all contributors) to the causal process of moving from the entertained premises to that conclusion; and many of these contributors will be bodily representations, either in their content or in their representational format.

That our minds are massively representation in this is, in my view, part of what embodied cognitive science reveals. Experimentalists can change the output of virtually any behavior by experimentally manipulating bodily processes. Why? Because those body-related inputs activate representations in distinctively bodily formats and which represent the same things as other perhaps nonbodily representations, and in doing so they disrupt or enhance the cognitive task at hand.

I illustrate with a parade result in the embodied literature, produced by Chen and Bargh (1999). In this series of experiments, Chen and Bargh instructed subjects to respond as quickly as possible to positively and negatively valenced words flashed on a computer monitor. In Experiment 1

Participants were randomly assigned to one of two experimental conditions. In the incongruent condition, the participants were instructed

to push the lever forward with their hand as quickly as possible when they judged the word as good and to pull the lever back toward them when they judged the word as bad. In the congruent condition, participants were given the opposite instructions: to pull the lever if the word was positive in meaning and to push the lever if the word was negative in meaning. (*ibid.*, 219)

It turns out that subjects respond significantly more quickly in the congruent condition than in the incongruent one. This has suggested various hypotheses, including the hypothesis that the meaning of evaluative words or the content of our evaluative concepts is embodied – the pulling of good things to us and the pushing away of bad things are parts of the very meanings of the relevant terms or concepts.

This kind of embodied semantics seems overblown to my mind, but perhaps there's a notion of meaning that's useful in this vicinity. Regardless, I take Chen and Bargh's results, and the many like them, to reveal something important about human cognitive processing and architecture; they support the massively representational view of the mind described above. One way that humans represent that something is good is by a certain kind of motor command to pull the thing closer (perhaps in conjunction with a representation of a positive-affect-related bodily state, the two of which have, as the result of experience, come to be associated). Thus, when the subject recognizes the word on the screen as have positive valence, this primes associated representations of goodness, including the motoric representation in question, which primes the subjects to react more quickly (that is, pull more quickly) in the congruent-good conditions.

None of this speaks much to the desire of many embodied theorists for a robustly first-personal theory of consciousness. In my view, this is because embodiment-oriented philosophers are misreading the embodied program. Embodiment-based experimental

results fit well into the massively representational framework, and what's more, that framework suggests a diagnosis: The embodiment-oriented theorist's hope (or even demand) for a first-person body-based theory of consciousness seems at least sometimes to rest on the assumption that there is a distinctive personal level (Varela 1996, 340).

Notice that nothing about Chen and Bargh's experiment rests on this assumption or compels its adoption. There are various ways of representing value – some of these are motor commands, some are encodings of the visual form of words. When more of these work together, we respond in a way that we associate with paying attention, being aware, or being conscious of a task – that is, we respond more readily and quickly. When fewer of these are active, and when countervailing ones are also activated, we respond less smoothly. All one has is the soup of representational units, their content and their causal relations – all operating at the level of what we would normally call the 'subpersonal'. But, calling them 'subpersonal' presupposes the existence of a personal level, which plays no role in accounting for the data. These experimental results as well as a wide range of others – including those of Lutz et al. discussed above! – are consistent with a completely "flat" psychological theory, according to which there is no distinction between supposed personal and subpersonal levels. And this view I recommend as an antidote to intuition-mongering about consciousness: all psychological and cognitive properties and processes are on the same ontological or scientific level. A mess of processes causally affect the behavior we associate with consciousness, attention, and awareness. And in embodied research, we probe and alter some of those processes by activating, in experimental conditions, some representations with bodily content or in a bodily format.

There's some irony to this discussion. Dennett provides large part of the inspiration for the embodied-cum-massively-representational-cum-flat picture of human psychology I propose. His discussion of a demon-like architecture (Dennett 1991) first exposed me to the general idea that loads of processes run in parallel and compete for control of the system's output (also see Calvin 1996). Yet, Dennett introduced the personal-subpersonal distinction! Be that as it may, so far as I can tell, the latter proposal – that there is a substantive personal-subpersonal distinction – is unmotivated in light of Dennett's views about consciousness. Dennett seems deeply wedded to the idea of personal level partly as a concomitant to his theory of the intentional stance (Dennett 1987), which is meant to provide everyday psychological explanation with its own integrity, independent of the nitty-gritty work of cognitive science. But, the commitment to a personal level seems to me to have little to do with the kind of observations about methodology and results in the cognitive sciences that drove Dennett's views about consciousness (Dennett 1988, 1991).

So, I recommending that we pry apart these two Dennettian theses – one concerning cognitive processing and architecture, the other concerning levels of reality or the layered nature of domains pertaining to mind. This prying apart allows for selective rejection, and here I'm rejecting the view of very many philosophers of cognitive science (see Drayson 2014, for a review and citations) who take personal-level phenomena to be given or to otherwise be established as the *explananda* of cognitive science, who take the detailed work of cognitive science to be the positing and testing of models of the subpersonal processes that implement personal-level capacities or personal-level dispositions. So far as I can tell, this introduces too many elements into our understanding of intelligent human behavior.

In closing, let me come at this from a slightly different angle. Tamar Gendler's work on what she calls aliefs (2008a and 2008b) also inspired some of what I say about the massively representational mind. To illustrate the force of aliefs, Gendler highlights cases of what she calls 'belief-discordant' behavior, in which something like the much-discussed System 1 processes (Kahneman 2011) – the alief-related processes – pull in one direction while the personal-level states pull in another. But, she also notes that lots of human behavior is belief-concordant; it is smooth and hitch-free partly because, roughly speaking, System 1 and System 2 processing points us in the same direction. It seems to me only a small and natural step from here to the flat view I recommend. Why think of the concordant cases as cases of agreement between levels, rather than as agreement between different processes operating at the same level?

VI. Conclusion

Sections 1 and 2 argued that none of the approaches to phenomenal consciousness that are dominant among embodied cognitive scientists solve the hard problem head on. Sections 3 and 4 argued attempts to integrate phenomenology – according to which refined introspection yields access to the bodily ground of consciousness – into embodied cognitive science add little of deep theoretical import to our scientific understanding of phenomenal consciousness; such a neurophenomenology is either a dualist concession to Chalmers, attaches an unjustified privilege to certain measurements over others, or constitutes nothing more than interesting everyday scientific work on the easy problems of consciousness. Thus, embodied cognitive science offers neither a head-on solution to the hard problem nor a distinctively embodied and phenomenological scientific methodology.

Section 5 offers an interpretation of embodied cognitive-scientific results, set within the framework of the thesis of a massively representational mind, that grounds an embodied jettisoning of the hard problem (Rupert 2011). But, it's one thing to eliminate the hard problem by deflating the intuition that conscious experiences have intrinsic qualitative characters, and it's another to offer a theory of conscious experience. Does the massively-representational-cum-flat view of cognition, inspired partly by work on embodiment, constitute a theory of phenomenal consciousness? To my mind that depends on whether the *explananda* associated with such terms as 'consciousness', 'attention', and 'awareness' find a home as natural kinds – that is, scientifically important kinds – in the models constituting our best cognitive science. Will a kind *consciousness* survive in these models? Or, will there be only such things as global workspaces, attentional circuits, working memory, and the cooperation of clusters of co-referring representations – the operations of none of which line up neatly with our standard usage of such terms as 'consciousness'? If our best models of cognition do without a rationally coherent personal-level – and include in its place only, say, a subpersonal narrative-constructing module that runs alongside all of the other so-called subpersonal mechanisms – does that alone eliminate consciousness from the human condition? In such cases, there's room for revision, elimination, and even socially driven stipulation (Churchland 1981, Stich 1996). At this stage in the empirical enterprise, there's no predicting the outcome with any confidence.

Works Cited

Ballard, D., Hayhoe, M., Pook, P., and Rao, R. (1997). Deictic Codes for the Embodiment of Cognition. *Behavioral and Brain Sciences* 20: 723–742.

- Barnier A. J., Sutton, J., Harris C. B., and Wilson, R. A. (2008). A Conceptual and Empirical Framework for the Social Distribution of Cognition: The Case of Memory. *Cognitive Systems Research* 9: 33–51.
- Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577–660
- Barsalou, L., Barbey, A. K., Simmons, W. K., and Santos, A. (2005). Embodiment in Religious Knowledge. *Journal of Cognition and Culture* 5: 14–57.
- Borrett, D., Kelly, S., and Kwan, H. (2000). Phenomenology, Dynamical Neural Networks and Brain Function. *Philosophical Psychology* 13: 213–28.
- Brooks, R. (1999). *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press.
- Calvin, W. (1996). *The Cerebral Code*. Cambridge, MA: MIT Press.
- Carruthers, P. (2011) *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford: Oxford University Press.
- Chalmers, D. (1995). Facing up to the Problem of Consciousness. *Journal of Consciousness Studies* 2: 200–219.
- Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.
- Chen, M., and Bargh, J. A. (1999). Consequences of Automatic Evaluation: Immediate Behavioral Predispositions to Approach or Avoid the Stimulus. *Personality and Social Psychology Bulletin* 25: 215–224.
- Churchland, P. M. (1981). Eliminative Materialism and the Propositional Attitudes. *Journal of Philosophy* 78: 67–90.
- Clark, A. (2003). *Natural-Born Cyborgs*. Oxford: Oxford University Press.
- Clark, A. (2008b). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.
- Colombetti, G. (2010). Enaction, Sense-Making, and Emotion. In Stewart et al. (2010), pp. 145–164.
- Dennett, D. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Dennett, D. (1988). Quining Qualia. In A. Marcel and E. Bisiach (Eds.), *Consciousness in Modern Science* (Oxford: Oxford University Press), pp. 42–77.

Dennett, D. (1991). *Consciousness Explained*. Boston: Little, Brown, and Company. 374–75.

Di Paolo, E. A., Rohde, M., De Jaegher, H. (2010). Horizons for the Enactive Mind: Values, Social Interaction, and Play. In Stewart et al. (2010), pp. 33–87.

Drayson, Z. (2014). The Personal/Subpersonal Distinction. *Philosophy Compass* 9: 338–346.

Gapenne, O. (2010). Kinesthesia and the Construction of Perceptual Objects. In Stewart et al. (2010), pp. 183–218.

Gallagher, S. (2005). *How the Body Shapes the Mind*. New York: Oxford University Press.

Gendler, T. (2008a). Alief and Belief. *Journal of Philosophy* 105: 634–663.

Gendler, T. (2008b). Alief in Action (and Reaction). *Mind & Language* 23: 552–585.

Gertler, B. (2011). *Self-Knowledge*. New York: Routledge.

Gibbs, R. (2006). *Embodiment and Cognitive Science*. Cambridge: Cambridge University Press.

Glenberg, A. (1997). What Memory is For. *Behavioral and Brain Sciences* 20: 1–19.

Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Help Us Think*. Cambridge: Belknap Press/Harvard University Press.

Harnad, S. (1990). The Symbol Grounding Problem. *Physica D* 42: 335–46.

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux

Kaschak, M., Madden, C., Therriault, D., Yaxley, R., Aveyard, M, Blanchard, A., and Zwaan, R. (2005). Perception of Motion Affects Language Processing. *Cognition* 94: B79–B89.

Lakoff, G., and Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Lakoff, G., and Nuñez, R. E. (2000). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. New York: Basic Books.

- Levine, J. (1983). Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly* 64: 354–361.
- Lutz, A. (2002). Toward a Neurophenomenology as an Account of Generative Passages: A First Empirical Case Study. *Phenomenology and the Cognitive Sciences* 1: 133–167.
- Lutz, A., Lachaux, J.-P., Martinerie, J., and Varela, F. J. (2002). Guiding the Study of Brain Dynamics by Using First-Person Data: Synchrony Patterns Correlate with Ongoing Conscious States during a Simple Task. *Proceedings of the National Academy of Sciences* 99: 1586–1591.
- Machery, E. (2009). *Doing without Concepts*. Oxford: Oxford University Press.
- Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.
- Petitot, J. (1995). Morphodynamics and Attractor Syntax: Constituency in Visual Perception and Cognitive Grammar. In R. F. Port and T. Van Gelder (eds.), *Mind as Motion* (Cambridge, MA: MIT Press, 1995) 227–281.
- Piccinini, G. (2010). How to Improve on Heterophenomenology: The Self-Measurement Methodology of First-Person Data. *Journal of Consciousness Studies* 17: 84–106.
- Prinz, J. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA: MIT Press.
- Prinz, J. (2012). *The Conscious Brain*. Oxford: Oxford University Press.
- Rupert, R. D. (1998). On the Relationship between Naturalistic Semantics and Individuation Criteria for Terms in a Language of Thought. *Synthese* 117: 95–131.
- Rupert, R. D. (2006). Review of Embodiment and Cognitive Science by Raymond Gibbs, in *Notre Dame Philosophical Reviews*, 2006.08.20.
- Rupert, R. D. (2009). *Cognitive Systems and the Extended Mind*. Oxford: Oxford University Press.
- Rupert, R. D. (2011). Embodiment, Consciousness, and the Massively Representational Mind. *Philosophical Topics* 39: 99–120.
- Rupert, R. D. (2013). Memory, Natural Kinds, and Cognitive Extension; or, Martians Don't Remember, and Cognitive Science Is Not about Cognition. *Review of Philosophy and Psychology* (special issue on memory and distributed cognition) 4: 25–47.
- Searle, J. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3: 417–24.

Shapiro, L. (2010). *Embodied Cognition*. New York: Routledge.

Stewart, J. (2010). Foundational Issues in Enaction as a Paradigm for Cognitive Science: From the Origin of Life to Consciousness and Writing. In Stewart et al. (2010), pp. 1–31.

Stewart, J., Gapenne, O., and Di Paolo, E. (eds.) (2010). *Enaction*. Cambridge, MA: MIT Press

Stich, S. (1996). *Deconstructing the Mind*. Oxford: Oxford University Press.

Tribble, E. B. (2005). Distributing Cognition in the Globe. *Shakespeare Quarterly* 56: 135–155.

Varela, F. (1996). Neurophenomenology: A Methodological Remedy for the Hard Problem. *Journal of Consciousness Studies* 3: 330–349.

Wheeler, M. (2005). *Reconstructing the Cognitive World: The Next Step*. Cambridge, MA: MIT Press.